AD-A267 180

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
28/2/93

3. REPORT TYPE AND DATES COVERED
Annual Technical Report 1/2/92-31/1/93

4. TITLE AND SUBTITLE
Cooperativity and 3-D Representation

5. FUNDING NUMBERS
AFOSR 91-0169
61102F
2313
AS

6. AUTHOR(S)
Patrick Cavanagh

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Harvard University
1350 Massachusetts Avenue
Holyoke Center
Cambridge, MA 02138-3826

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research / NL
Building 410, Room c124
Bolling AFB, DC 20332-6448
Dr. Tangney

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

DTIC ELECTE JUL 26 1993 S E D

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

93-16688

13. ABSTRACT (Maximum 200 words)

Recent evidence indicates that the early stages in visual processing may be broken into several parallel streams that are specialized for the analysis of different visual attributes. A contour localization task showed that all attributes can contribute equally to border localization — no particular attribute dominated position decisions. A series of experiments on transparency perception showed that transparency is analyzed rapidly (within 60 msec) and influences early levels of visual processing. We have also investigated the early stages that lead from the initial 2-D representation to object recognition. Visual priming studies have been completed which suggest that object recognition begins, not with the construction of a 3-D model, but with a crude match of 2-D views to internal prototypes. The prototype that has the best match then guides the construction of an internal 3-D model. An analysis of drawing techniques has led to ideas about memory codes for the depth structure of an image, to simplified mechanisms for understanding shadows and shading and to renewed interest in "isophot" models of shading. Visual search studies have shown that some scene features may be rapidly suppressed. For example, shadows appear to be identified early and discounted in order to allow object contours to be processed. Finally, long-term practice in visual search tasks leads to learning of both object-centered and retinotopic properties of the stimuli.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT

2 5 JUN 1993

Cooperativity and 3-D Representation

Objectives

Our work has concentrated on how 2-D information is built up from the parallel analysis of a set of visual attributes and how this information contacts memory in order to construct 3-D representations of the visual scene. We are interested in the coding of image contours and how they arise from the various attributes which can define contours. We are examining the decomposition of image values into object features (reflectance, orientation, 3D position) and illumination features (shadows, shading, highlights) and especially how the perception of transparency leads to the distribution of image values to two or more superimposed surfaces. Finally, we are studying the initial contact between the image contours and memory in recognition.

Progress Report

Cross-media cooperation in contour localization. Josée Rivest's thesis research, described in last spring's annual report, was completed in the fall. Her final experiments measured the interactions between contours defined by different attributes. She found that all contours interacted in a similar manner suggesting a common final representation for contours independent of the attribute that defines them. For example, a color contour strongly attracts an adjacent luminance contour, displacing its perceived position towards the color contour. The reverse is also true, a luminance contour attracts an adjacent color contour. These results challenge the models of Gregory and Grossberg that predict that luminance is a priviledged signal for contours. This work was presented at ARVO in 1992 and manuscripts are being prepared for publication. Josée Rivest will return to ouer lab this summer to work on these manuscripts and two additional experiments.

Object recognition: priming. In our model, recognition starts with an initial, crude 2-D match that selects a "best" prototype to explain the image data. This is followed by more sophisticated 3-D analyses to complete the recognition process. Our first experiment showed a priming effect of contours in recognition even though the contours alone were uniformative for the task. This priming is probing an early (about 100 msec into processing) 2D stage of recognition. Takeo Watanabe will be returning for the summer to continue work on this project.

The art of perceiving contours. I have been examining art and the origins of various drawing techniques to discover what artists know about vision that scientists have not yet discovered. This has led in three different directions.

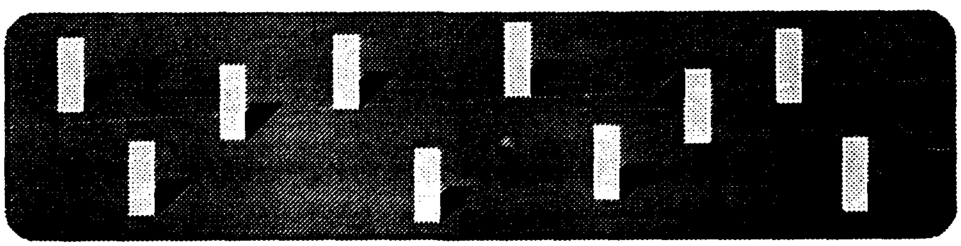
The first follows from the surprising effectiveness of line drawings. This earliest form of art (20,000 BC) uses lines to represent depth discontinuities. Illumination discontinuities (shadow boundaries) and material boundaries cannot be depicted as lines without leading to serious ambiguities in these drawings. Why would lines work at all to support the perception of depth discontinuities?? There are no lines in the real world that surround objects or mark occluding borders, there are edges — discontinuities in brightness, color or texture. The effectiveness of lines reveals a critical aspect of internal coding for depth discontinuities. An object may differ from its background in any or all surface attributes — color, brightness, motion or texture — so it is reasonable that some attribute-invariant representation of depth discontinuities should evolve. I would like to call this the 'depth sketch'. Apparently, the mechanisms that achieve this attribute invariance also respond to lines. This may be relevant to compact coding for memory of

objects indepenent of direction of lighting. I would like to follow this line of reasoning with several experiments during the coming year.

Second is the independent evolution of the techniques for representing shadows (dating from about 400 BC) and those for representing shading (dating from about 1400 BC). Once both were present in paintings, it was evident that no concept of a consistent light source was used or required. Shadows and shading were often depicted as generated from different directions of illumination. Nevertheless the impression of depth from both remains quite convincing. This demonstrates that human analysis of these factors does not depend on any hypothesis of consistent or even realizable lighting. Each seems to contribute a local cue to depth which is interpretable on its own. As described below, Ron Rensink and I are pursuing this simplified lighting analysis with visual search tasks.

Third, I was intrigued with the field of electron microscopy because in this area, the richly 3D images of pollen bits or whatever are always present in negative contrast. Earlier work in our laboratory had shown how vulnerable depth from shadows was to reversed contrast — the 3D interpretation was destroyed. That is not true for shading, however, and electron microscopy has the property that the direction of the illumination is the same as the direction of view so that there are no shadows anywhere in the image, only shading. Note that no shape-from-shading algorithm which tries to extract the direction of the illuminant can analyze these negative images because there is no realizable light source or sources which can produce the image (only "black" light can do so). Nevertheless, we see the depth immediately. This implies that our internal representation is robust to reversals of contrast. One such representation is the isophots of the images, the lines of equal luminance. They capture the flow field of the brightness gradient and have the same structure when the contrast is reversed. I am interested in following up this observation with work on the internal coding for shading and its underlying physiological mechnisms as well as the relation of this internal code to the depth perception seen for surfaces of ruled lines (as in Kent Stevens' work). These ruled line surfaces (which have few or no equivalents in the natural world) may activate the same analyses as shaded surfaces and so the depth that they evoke may be another example of art exploiting the fortuitous consequences of the internal code.

Object features and scene attributes in visual search. Ron Rensink and I have shown that preattentive vision is also sensitive to scene structure defined by shadows.

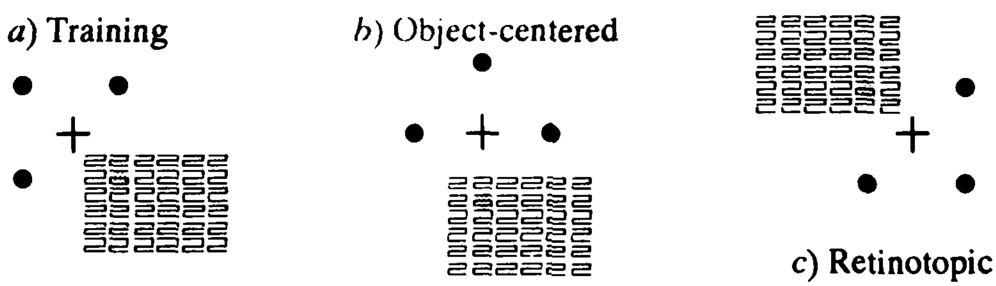


Evidence for this is based on visual search experiments involving simple black-and-white figures that correspond to shadowed posts. Search is rapid when the target has an anomalous shadow (view figure from rightside up), but slows down when the distractors have anomalous shadows (view figure from upside down). This asymmetry in speed is destroyed when the areas corresponding to shadows are made lighter than the background or are surrounded by a thin white line; for these conditions, search is evidently based on image properties alone. This rapid

We believe that our results with shadows is just one example of feature suppression by the visual system. We know that shadows must be rapidly identified and suppressed for object contours to be processed. We predict that other image features like the brightness patterns of highlights may suffer a similar fate. That is, they are identified as features of the scene lighting and become detached from the object and discounted or suppressed. They leave residual effects in the interpretation of the object's surface material but are not easily accessed to determine their actual brightness or shape. Additional experiments already underway support this possibility.

Finally, our interest in the internal representation of shading as an oriented field of isophots (or at least short oriented segments) is still at an exploratory stage. We will digitize live scenes in our laboratory using a camera with the lighting fixed to camera lens. This lighting produces only shading in the image. The only shadows arise from secondary reflection which can be minimized with dark walls and materials. The images will be viewed in positiove, negative and in isophot displays. The isophot display is produced by sending the image intensity levels through a lookup table which assigns white and black alternatively to the ascending intensity levels. This creates a contour structure that follows the isophots of the image. We believe that all three versions will be similarly interpretable. The important advance over previous attempts to look at images in this fashion is that we will eliminate shadows and thus the major cause of difficulties in interpretation. We will also test whether continuous isophots are required for the depth to be recovered or whether oriented fields of random linesegments is sufficient.

Learning and visual search. Satoru Suzuki and I examined whether pattern discrimination learning is specific to the location of the target in retinal coordinates or to the location of the target within the array of distractors. The first experiment was a threshold task and the second was a reaction time task. In the training phase of the first experiment, the target, , was placed at location (2,2) in the upper left corner of a 6x6 rectangular array of 's and the array was positioned such that the fell randomly on one of the four corners of a virtual square centered at the fixation point (a below). The task was to decide if the flashed array contained the target. Presentation duration was varied between 50 and 250 msec. Following 20 days of 2400 trials per day, threshold durations for 75% correct responses dropped from 190 msec to 120 msec. The specificity of this learning was tested by measuring thresholds on modified stimulus arrays: retinotopic specificity was tested by rotating the virtual square of the four locations of the target item by 45° (b below), and object-centered specificity by moving the target item into the lower right corner of the array (c below). The results showed that learning was both retinotopic and object-centered. The learning persisted undiminished over four weeks. In contrast, extended practice on an analogous reaction time task using identical stimuli suggested that learning for this task was object-centered but not retinotopic and also rather short term (reaching plateau/decaying within hours). A manuscript is in preparation.



Publications during grant period (* indicates support from grant)

Arguin, M., Cavanagh, P., & Joanette, Y. (in press). Visual feature integration with an attentional deficit. *Brain and Cognition* .

Arguin, M., Joanette, Y., & Cavanagh, P. (in press). Visual search for feature and conjunction targets with an attention deficit. *Journal of Cognitive Neuroscience*.

* Cavanagh, P. (1993). The perception of form and motion. *Current Opinion in Neurobiology*, **3**, 177-182.

* Watanabe, T., & Cavanagh, P. (in press). Surface decomposition in subjective transparency. *Spatial Vision* .

* Watanabe, T., & Cavanagh, P. (in press). Implicit X junctions. *Vision Research*..

Cavanagh, P. (1992). Attention-based motion perception. *Science*, **257**, 1563-1565.

Cavanagh, P., Adelson, E. H., & Heard, P. (1992). Vision with equiluminant colour contrast: 2. A large-scale technique and observations. *Perception*, **21**, 219-226.

Jolicoeur, P., & Cavanagh, P. (1992). Mental rotation, physical rotation, and surface media. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 371-384.

* Shioiri, S., & Cavanagh, P. (1992). Visual persistence of figures defined by relative motion. *Vision Research*, **32**, 943-951.

Shioiri, S., & Cavanagh, P. (1992). Achromatic form perception is based on luminance not brightness. *Journal of the Optical Society of America A* **9**, 1672-1681.

* Takeichi, T., Shimojo, S., & Watanabe, T. (1992). Neon flank and illusory contour: interaction between the two processes leads to color filling-in. *Perception*, **21**, 313-324.

* Takeichi, T., Watanabe, T., & Shimojo, S. (1992). Illusory occluding contours and surface formation by depth propagation. *Perception*, **21**, 177-184.

* Watanabe, T., & Cavanagh, P. (1992). Depth capture and transparency of regions bounded by illusory, chromatic, and texture contours. *Vision Research*, **32**, 527-532.

* Watanabe, T., & Cavanagh, P. (1992).The role of transparency in perceptual grouping and pattern recognition. *Perception*, **21**, 133-139.

* Watanabe, T., Zimmerman, L. & Cavanagh, P. (1992).Orientation-contingent color aftereffects mediated by subjective transparent surfaces. *Perception & Psychophysics*, **52**, 161-166.

Bruck, M., Cavanagh, P., & Ceci, S. J. (1991). Fortysomething: Recognizing faces at one's 25th reunion. *Memory & Cognition*, **19**, 221-228.

Cavanagh, P. (1991). Short-range vs long-range motion: not a valid distinction. *Spatial Vision*, **5**, 303-309.

Cavanagh, P. (1991). The contribution of color to motion. In A. Valberg and B. B. Lee (eds.) *From pigments to perception*. New York: Plenum, 151-164.

* Cavanagh, P. (1991). What's up in top-down processing? In A. Gorea (ed.) *Representations of Vision: Trends and Tacit Assumptions in Vision Research*, Cambridge, UK: Cambridge University Press, 295-304.

Cavanagh, P. (1991). Vision at equiluminance. In J. J. Kulikowski, I. J. Murray, and V. Walsh (eds.) *Vision and Visual Dysfunction Volume V: Limits of Vision*. Boca Raton, FL: CRC Press, 234-250.

Cavanagh, P., & Anstis, S. M. (1991). The contribution of color to motion in normal and color-deficient observers *Vision Research*, **31**, 2109-2148.

Tyler, C. W., & Cavanagh, P. (1991). Purely chromatic perception of motion in depth: Two eyes as sensitive as one. *Perception & Psychophysics*, **49**, 53-61.

* Watanabe, T., & Cavanagh, P. (1991). Texture and motion spreading, the aperture problem, and transparency. *Perception & Psychophysics*, 50, 459-464.

Participating Professionals

Personnel on the grant this past year were myself (80% summer salary), Ron Rensink (Postdoctoral Fellow), and Satoru Suzuki (graduate student and summer research assistant). Raynald Comtois, our Senior Systems Analyst (25% salary), has been funded on this grant since the supplemental award last year. Dr. Josée Rivest (graduate student) successfully defended her thesis last September and is now an assistant professor at the Glendon Campus of York University in Toronto, Ontario. Dr. Takeo Watanabe (Research Associate) has been an associate professor at State University of Arizona West since last August

Interactions, conference papers during grant period (* support from grant)

* Cavanagh, P. (1993). The art of perceiving contours. The Clarence H. Graham Memorial Lecture to the EPA, Washington DC, April.

Cavanagh, P. (1993). There then but here now. Paper presented to the Conference on Linking Psychophysics, Neurophysiology, and Computational Vision, Rutgers, April.

Culham, J., & Cavanagh, P. (1993). Capture of luminance stimuli by moving equiluminous color gratings and by attentive tracking. *Investigative Ophthalmology and Visual Science*, **34**, 976.

* Rensink, R., & Cavanagh, P. (1993). Processing of shadows at preattentive levels. *Investigative Ophthalmology and Visual Science*, **34**, 1288.

* Suzuki, S., & Cavanagh, P. (1993). Retinotopic vs object-centered specificity of learning differs for threshold and reaction time tasks. *Investigative Ophthalmology and Visual Science*, **34**, 1233.

Wang, Q., & Cavanagh, P. (1993). Acquiring familiarity effects in visual search with Chinese characters. *Investigative Ophthalmology and Visual Science*, **34**, 1236.

* Cavanagh, P. (1992). What are the orthogonal dimensions of vision? Dissection vision with perceptual studies in normals. Paper presented to the CVS Symposium on Localizing Visual Function in the Brain, Rochester, June.

Intriligator, J., & Cavanagh, P. (1992). An object-specific attentional facilitation that does not spread to adjacent locations. *Investigative Ophthalmology and Visual Science*, **33**, 1263.

* Rivest, J., & Cavanagh, P. (1992). Interattribute interactions in contour localization. *Investigative Ophthalmology and Visual Science*, **33**, 1342.

* Suzuki, S., & Cavanagh, P. (1992). Facial expression as an emergent feature in visual search. *Investigative Ophthalmology and Visual Science*, **33**, 1355.

Wang, Q., & Cavanagh, P. (1992). Familiarity and pop-out in visual search. *Investigative Ophthalmology and Visual Science*, **33**, 1262.

* Watanabe, T., & Cavanagh, P. (1992). Transparent surfaces defined by implicit X junctions. *Investigative Ophthalmology and Visual Science*, **33**, 706.

Cavanagh, P. (1991). No slowing for active motion perception at equiluminance. *Investigative Ophthalmology and Visual Science*, **32**, 894.

Charles, E., Logothetis, N. K., & Cavanagh, P. (1991) The response of cells in area MT to stimulation of blue-sensitive cones. *Society for Neuroscience Abstracts*, **17**, 440.

Intriligator, J., Nakayama, K., & Cavanagh, P. (1991). Attention tracking of multiple moving objects at different scales. *Investigative Ophthalmology and Visual Science*, **32**, 1040.

* Rivest, J., & Cavanagh, P. (1991). Cross media cooperation in localization of contours. *Investigative Ophthalmology and Visual Science*, **32**, 1024.

* Takeichi, T., Shimojo, S., & Watanabe, T. (1991). Neon flank and illusory contours: monocular and binocular processes integrated in the color filling-in. *Investigative Ophthalmology and Visual Science*, **32**, 696.
* Watanabe, T., & Cavanagh, P. (1991). Transparency facilitates recognition of overlapping stimuli even for very short exposures. *Investigative Ophthalmology and Visual Science*, **32**, 1024.